Proc. Natl. Acad. Sci. USA
Vol. 92, pp. 7764–7768, August 1995
Biochemistry

Proenzyme of *Manduca sexta* phenol oxidase: Purification, activation, substrate specificity of the active enzyme, and molecular cloning

(hawkmoth/melanization/sclerotization/insect immunity)

MARTIN HALL*†, TIMOTHY SCOTT‡, MANICKAM SUGUMARAN‡, KENNETH SÖDERHÄLL†, AND JOHN H. LAW*§

*The Department of Biochemistry and Center for Insect Sciences, University of Arizona, Tucson, AZ 85721; ‡The Department of Biology, University of Massachusetts, Boston, MA 02125; and †The Department of Physiological Botany, University of Uppsala, Villavägen 6, 752 36 Uppsala, Sweden

Contributed by John H. Law, April 21, 1995

ABSTRACT **Phenol oxidase (PO) was isolated as a proenzyme (pro-phenol oxidase, pro-PO) from the hemolymph of *Manduca sexta* larvae and purified to homogeneity. Pro-PO exhibits a M_r of 130,000 on gel filtration and two bands with an apparent M_r of ≈100,000 on SDS/PAGE, as well as size-exclusion HPLC. Activation of pro-PO was achieved either by specific proteolysis by a cuticular protease or by the detergent cetylpyridinium chloride at a concentration below the critical micellar concentration. A cDNA clone for *M. sexta* pro-PO was obtained from a larval hemocyte cDNA library. The clone encodes a polypeptide of ≈80,000 Da that contains two copper-binding sites and shows high sequence similarity to POs, hemocyanins, and storage proteins of arthropods. The *M. sexta* pro-PO, together with other arthropod pro-POs, contains a short stretch of amino acids with sequence similarity to the thiol ester region of α-macroglobulins and complement proteins C3 and C4.**

Melanin biosynthesis occurs widely, not only in animals but also in plants and fungi (1). Phenol oxidase (PO), which possesses both tyrosinase activity (monophenol monooxygenase; monophenol, L-dopa:oxygen oxidoreductase; EC 1.14.18.1), as well as *o*-diphenol oxidase (1,2-benzenediol:oxygen oxidoreductase; EC 1.10.3.1), is responsible for initiating the biosynthesis of melanin (2, 3). In arthropods and especially in insects, PO is uniquely involved in another important biochemical process—cuticular sclerotization (hardening). Sclerotization is vital for the survival of all insects, as it affords protection to the soft invertebrate body (3–8). During sclerotization, PO-generated quinones participate in the quinone tanning process or serve as substrates for quinone isomerases that convert quinones to quinone methides for quinone methide sclerotization (9–13). Certain quinone methides are converted by another cuticular enzyme, quinone methide isomerase, to 1,2-dehydro-*N*-acyldopamine derivatives (14). These compounds are further oxidized by PO to reactive quinone methide imine amides (15, 16), which serve as the reactive intermediates of α,β-sclerotization. Quinones, quinone methides, and quinone methide imine amides react with cuticular proteins and chitin, forming eventually cross-linked cuticle.

Apart from participating in sclerotization and melanization of cuticle, POs are also known to be involved in two other physiologically important processes—defense reactions (arthropod immunity) and wound healing. During invasion by a foreign organism, pro-phenol oxidase (pro-PO) present in the hemocytes is released and activated to produce melanin pigments for deposition on the intruder (17–19). The damage that can be caused by the foreign organisms is thus limited by encapsulation and melanization. Similarly during wounding of insect cuticle, PO causes massive deposition of melanin pigment at the wound site to prevent hemolymph loss and to block the entry of opportunistically invading microorganisms (20).

Given such important functions, it could be expected that PO would be one of the best studied of insect enzymes. Surprisingly, in spite of the numerous investigations carried out on this enzyme, only limited knowledge is available. Studies on PO are severely hampered by its instability, "stickiness," loss of activity during purification, insolubilization, and inactivation during reaction. Consequently, very few researchers have succeeded in examining the detailed biochemical properties of this enzyme. However, the inactive zymogen, pro-PO, from *Bombyx mori* (21), *Manduca sexta* (22), *Hyalophora cecropia* (23), *Calliphora erythrocephala* (24), *Musca domestica* (25), *Blaberus discoidalis* (26), and *Tenebrio molitor* (27), as well as from the crayfish *Pacifastacus leniusculus* (28), has been purified to or near homogeneity. However, knowledge of the structure, properties, and regulation of this enzyme that is needed for an understanding of its critical role in insect biochemistry and physiology is still lacking. As a contribution in this direction, we have achieved the purification, characterization, and molecular cloning of pro-PO from *M. sexta* and present the results in this paper.¶

MATERIALS AND METHODS

Animals. The eggs of *M. sexta* were provided by J. S. Buckner of the Metabolism and Radiation Research Laboratories, U.S. Department of Agriculture, Fargo, ND. Eggs were hatched, and the larvae were raised on a synthetic diet (29).

Enzyme Purification. All operations were done at 0–5°C, unless otherwise stated. Centrifugations were performed at 12,000 × *g* for 10 min. Thirty *M. sexta* larvae (fifth instar) were bled into 50 ml of 30% (vol/vol) saturated ammonium sulfate solution. After adjusting the final volume of the solution to 80 ml and the ammonium sulfate concentration to 30% saturation, the contents were centrifuged. The supernatant was brought to 60% saturation with respect to ammonium sulfate, and proteins precipitated within 30 min were collected by centrifugation. The pellet was dissolved in a minimum amount of 10% saturated ammonium sulfate/10 mM sodium phosphate buffer, pH 6.0 (buffer A) and loaded on a phenyl-Sepharose column (4.5 × 2.5 cm) pre-equilibrated with the same buffer. The column was eluted with buffer A for 10 min, followed by a linear gradient (descending) of buffer A and water for the next 20 min and finally with water for 20 min. A flow rate of 2 ml/min was maintained. Fractions that contained pro-PO were pooled, dialyzed against water (4 liters,

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: CPC, cetylpyridinium chloride; PO, phenol oxidase; pro-PO, pro-phenol oxidase.

§To whom reprint requests should be addressed.

¶The sequence reported in this paper has been deposited in the GenBank data base (accession no. L42556).

three changes), and chromatographed on a DEAE-Sepharose column (9 × 2.5 cm) equilibrated with water. After loading the sample, the column was washed with water for 20 min followed by a 40-min linear gradient of water and 100 mM sodium phosphate buffer, pH 5.0 (buffer B). Final washing was done with buffer B for 30 min at a flow rate of 2 ml/min. The pro-PO solution obtained was placed in a dialysis membrane, which was packed in solid sucrose to concentrate the solution. The concentrated solution was chromatographed on a Sephacryl HR-S-300 column (45 × 2.5 cm) equilibrated with 50 mM sodium phosphate buffer, pH 6.0. Fractions containing pro-PO were pooled and used as pure enzyme.

Assay of Pro-PO. A reaction mixture (1 ml) that contained 2 mM dopamine, 50 mM sodium phosphate buffer, pH 6.5, and enzyme protein was incubated at room temperature, and the absorbance increase at 475 nm associated with the oxidation of substrate was continuously monitored in a spectrophotometer after addition of 10 μl of 10% cetylpyridinium chloride (CPC) to activate the pro-PO. For some experiments, the assay was conducted on a Yellow Springs Instruments model 53 oxygen electrode using the standard assay mixture at 30°C. Cuticular protease known to activate pro-PO was isolated as described by Aso *et al.* (22).

Protein Sequence Determination. Edman degradation of purified pro-PO yielded no phenylthiohydantoin derivative of any amino acid, suggesting a blocked amino terminus. Therefore, pro-PO was proteolytically digested. Purified enzyme was reduced with dithiothreitol, alkylated, and digested with trypsin. The resultant peptides were separated on a reversed-phase C_{18} HPLC column. Amino acid sequences of two peptides were determined by using an Applied Biosystems model 477A instrument equipped with on-line phenylthiohydantoin analysis.

cDNA Cloning and Sequence. A cDNA library from hemocytes of fifth-instar larvae of *M. sexta* was provided by Michael R. Kanost, Kansas State University. The cDNA had been synthesized using an oligo(dT) primer and then directionally cloned (*Eco*RI to *Xho* I, 5′ → 3′) into Uni-ZAP vector (Stratagene).

Initially a 700-bp PCR product coding for the 3′ end of the cDNA for pro-PO was obtained by serendipity while searching for insect proteins related to α-macroglobulins and complement proteins. A degenerate oligonucleotide primer, GGIT-GYGGIGARCARAAYATG (21 nt, I = inosine), was designed to correspond to the amino acid sequence GCGEQNM, which is conserved in α-macroglobulin and complement proteins C3 and C4 from vertebrates (30), as well as in α-macroglobulin-like proteins of invertebrates (31–34). In these proteins, the cysteine and glutamine participate in the formation of a reactive thiol ester group. PCR amplification of the hemocyte library with this primer and a vector (T7)-specific primer yielded a product of ≈700 bp, which was cloned into the pBluescript vector, using a TA cloning kit (Invitrogen) and sequenced by the dideoxynucleotide chain-termination method (Sequenase 2.0, United States Biochemical). The PCR product contained the sequence of the degenerate primer and an open reading frame of ≈300 nt, followed by a stop codon and 300 nt of untranslated DNA [including a polyadenylylation signal and a poly(A) tail] and finally some vector sequence and the T7 primer. The translated sequence showed similarity to arthropod hemocyanins and to some insect larval storage proteins, as well as to the crayfish pro-PO (35). The cloned PCR product was excised from the plasmid, isolated, labeled with ^{32}P (random primer labeling kit, GIBCO/BRL) and used to screen the hemocyte library. From a total of 210,000 plaques, 10 positive plaques were selected and rescreened twice, after which single plaques were selected for sequencing. The 10 clones were sequenced with M13 forward and reverse primers. These clones proved to be of different lengths (Fig. 1), but all contained the sequence of the PCR product at the 3′ end. The 5′ ends were different in all 10 clones, and they were found to be partially overlapping—the end of the interpretable sequence on one clone was identical to the beginning of another, etc. In this way the sequence of the longest clone could be determined for nearly its entire length. Finally, the entire sequence of the longest clone, PO7, was determined in both directions by designing specific sequencing primers ≈200 nt apart (Fig. 1).

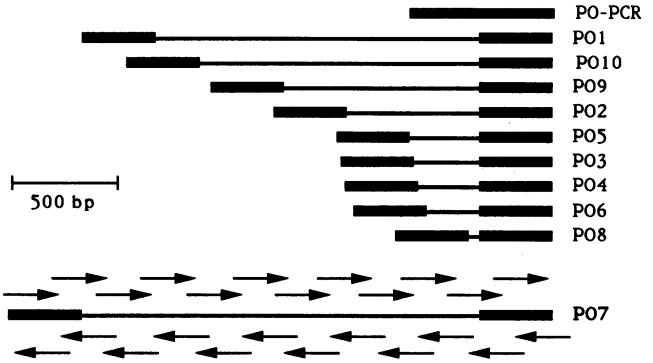

FIG. 1. Sequencing strategy for *M. sexta* pro-PO. The 700-bp PCR product (PO-PCR) was used to screen a hemocyte cDNA library. Ten positive clones (PO1–10) were selected for sequencing. Areas marked in thicker lines show sequences obtained with vector-specific primers from either 3′ or 5′ ends. Suitable sequence primers, ≈200 nt apart, were designed from the determined sequences and used for sequencing both strands of the longest clone, PO7, as indicated by arrows.

RESULTS AND DISCUSSION

Purification of Pro-PO. With ammonium sulfate precipitation, phenyl-Sepharose chromatography, DEAE-Sepharose chromatography, and Sephacryl S-300 chromatography, pro-PO was purified ≈2000-fold. The homogeneous preparation of pro-PO, when subjected to SDS/PAGE followed by silver staining, revealed two closely spaced bands and no other protein bands (Fig. 2*A*). Both these bands were also stained positively for PO activity after activation by CPC (Fig. 2*B*).

Properties of Pro-PO. The approximate M_r of pro-PO as estimated by chromatography on Sephacryl S-300 is 130,000 (data not shown). On SDS/PAGE, two bands corresponded to

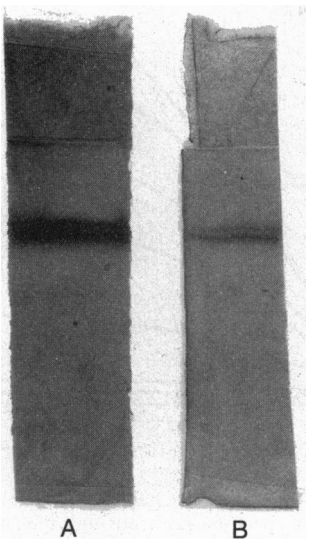

FIG. 2. SDS/PAGE of purified pro-PO and PO. (*A*) Stained with silver. (*B*) Stained with 4-methylcatechol followed by 3-methyl-2-benzothiazolinone hydrazone. (Note the presence of two closely migrating bands on each gel.)

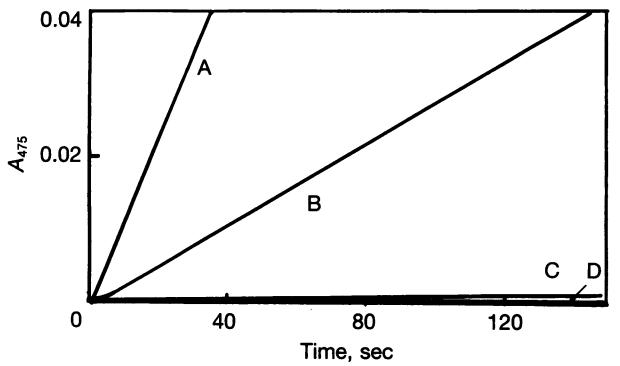

FIG. 3. Activation of pro-PO by different reagents. A reaction mixture (1 ml) containing 2 mM dopamine, enzyme protein (2 μg), and 0.02% detergent in 50 mM sodium phosphate buffer, pH 6.5, was incubated, and the increase in absorbance at 475 nm was continuously monitored. Detergents used were CPC (*A*) and SDS, sodium caprylate, deoxycholate, sodium *N*-laurylsarcosidate, digitonin, Triton X-100, CHAPS, Tween 20, Nonidet P-40, Brij 58 or octyl β-D-glucopyranoside (*D*). Protease treatment was done at pH 8.0, and then the enzyme activity was assayed at pH 6.0. A reaction mixture (1 ml) containing enzyme protein, protease (2–10 units), in ammonium bicarbonate buffer, pH 8.0, was incubated for 10 min, and a 100-μl aliquot was used for assaying pro-PO activity by using a standard assay. (*A*) Control with CPC without protease. (*B*) *M. sexta* cuticular protease. (*C*) Trypsin, chymotrypsin, or subtilisin.

98,000 and 100,000 Da (Fig. 2*B*). A M_r of 90,000 was estimated by HPLC gel-permeation analysis (data not shown). On the basis of the deduced amino acid sequence, the pro-PO should have a M_r of ≈80,000. At present why the protein exhibits a higher M_r of 130,000 on a Sephacryl S-300 column is unclear.

Activation of Pro-PO. Pro-PO of insects can be activated by two different mechanisms, one involving the proteolytic cleavage of the proenzyme and the other involving detergent activation (17). The purified *M. sexta* pro-PO can be activated by both these modes as shown in Fig. 3. The cuticular protease isolated from the same organism readily activated the proenzyme. Of the various detergents tested, only the cationic detergent CPC activated the *M. sexta* proenzyme. The zwitterionic detergent 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate (CHAPS) did not activate the pro-PO. Similarly, neither anionic detergents, such as SDS, sodium caprylate, deoxycholate, sodium laurylsarcosidate, nor nonionic detergents such as octyl glucopyranoside, Nonidet P-40, Triton X-100, digitonin, Brij 58, and Tween 20 activated the proenzyme (Fig. 3). The activation caused by CPC (K_a = 0.046 M) occurred well below its critical micellar concentration (0.8 M).

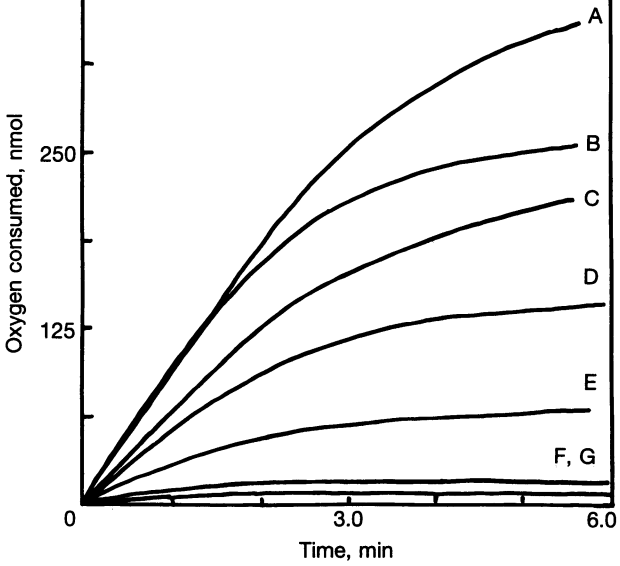

FIG. 4. Substrate specificity of hemolymph PO. A reaction mixture containing 1 mM of specified substrate, enzyme, and CPC in 1 ml of 50 mM sodium phosphate buffer, pH 6.5, was incubated at 30°C, and the oxygen uptake associated with the reaction was continuously monitored. Substrates used were *N*-β-alanyldopamine (*A*), 4-methylcatechol (*B*), dopamine (*C*), *N*-acetyldopamine (*D*), catechol (*E*), dopa (*F*), and *N*-acetylnonepinephrine (*G*).

Substrate Specificity of PO. The substrate specificity of CPC-activated pro-PO is shown in Fig. 4. *N*-β-Alanyldopamine and 4-methylcatechol proved superior substrates, and dopa, often used as a substrate for routine assay of insect PO, proved to be a poor substrate for *M. sexta* hemolymph PO. In the light of these results, routine use of dopa as a substrate for the insect enzyme is not recommended.

cDNA Sequence. The sequence of cDNA encoding *M. sexta* pro-PO, along with the deduced amino acid sequence, is shown in Fig. 5. It contains 2466 nt, including a polyadenylylation signal and poly(A) tail, which are underlined. The ATG encod-

```
   1  GGAACTGTGAAGTGGGTATACTAAAAAGTTTTTTTTTATTTATTTAAAAATACGTTTAT   60
  61  CAGACACCATGGCTGATATTTTTGATAGTTTCGAACTCCTCTACGACCGTCCAGGAGAAC  120
          M  A  D  I  F  D  S  F  E  L  L  Y  D  R  P  G  E  P
 121  CTATGATTAACACCAAAGGGGAAGACAAAGTTCTATTCGAACTCACTGAACAATTTCTGA  180
      M  I  N  T  K  G  E  D  K  V  L  F  E  L  T  E  Q  F  L  T
 181  CCCCGGAATACGCCAACAATGGTCTGGAGTTGAATAACCGCTTCGGTGATGAGGAGGAGG  240
      P  E  Y  A  N  N  G  L  E  L  N  N  R  F  G  D  E  E  E  V
 241  TGTCTCGTAAAATAATATTGAAGAATCTTGACAAGATTCCAGAGTTTCCTAAAGCTAAAC  300
      S  R  K  I  I  L  K  N  L  D  K  I  P  E  F  P  K  A  K  Q
 301  AACTCCCAAACGATGCCGATTTCTCCCTTTTCCTGCCCAGCCATCAAGAAATGGCTAATG  360
      L  P  N  D  A  D  F  S  L  F  L  P  S  H  Q  E  M  A  N  E
 361  AAGTCATTGATGTCCTAATGAGTGTAACTGAGAACCAACTACAAGAACTCCTGTCTACTT  420
      V  I  D  V  L  M  S  V  T  E  N  Q  L  Q  E  L  L  S  T  C
 421  GTGTGTATGCCCGAATCAATCTCAACCCGCAGCTGTTCAACTACTGCTACACTGTTGCCA  480
      V  Y  A  R  I  N  L  N  P  Q  L  F  N  Y  C  Y  T  V  A  I
 481  TTATGCACAGACGTGACACGGGTAAGGTTCGTGTACAGAACTATGCAGAAATTTTCCCTG  540
      M  H  R  R  D  T  G  K  V  R  V  Q  N  Y  A  E  I  F  P  A
 541  CAAAGTTTTTGGACTCTCAAGTATTCACCCAGGCCCGTGAAGCTGCAGCAGTCATCCCGA  600
      K  F  L  D  S  Q  V  F  T  Q  A  R  E  A  A  A  V  I  P  K
 601  AAACTATTCCTCGAACTCCTATCATCATTCCACGAGACTACACTGCTACCGACTTGGAAG  660
      T  I  P  R  T  P  I  I  I  P  R  D  Y  T  A  T  D  L  E  E
 661  AAGAGCATCGCCTGGCGTACTGGCGTGAAGACCTCGGTATCAATCTTCACCATTGGCATT  720
      E  H  R  L  A  Y  W  R  E  D  L  G  I  N  L  H  H  W  H  W
 721  GGCATCTCGTTTACCCGTTCTCTGCTAGCGACGAGAAAATTGTAGCCAAAGACCGTCGCG  780
      H  L  V  Y  P  F  S  A  S  D  E  K  I  V  A  K  D  R  R  G
 781  GGGAGCTGTTCTTCTACATGCACCAGCAAATCATTGCCAGATACAACTGCGAGCGTTTGT  840
      E  L  F  F  Y  M  H  Q  Q  I  I  A  R  Y  N  C  E  R  L  C
 841  GCAATTCTCTGAAGAGGGTAAAGAAATTCAGCGACTGGCGCGAGCCTATCCCGGAGGCAT  900
      N  S  L  K  R  V  K  K  F  S  D  W  R  E  P  I  P  E  A  Y
 901  ATTACCCTAAACTAGACAGCTTGACTTCAGCCCGCGGCTGGCCGCCACGCCAAGCTGGTA  960
      Y  P  K  L  D  S  L  T  S  A  R  G  W  P  P  R  Q  A  G  M
 961  TGCGTTGGCAAGACCTGAAGAGACCTGTGGACGGTCTAAACGTTACAATTGATGACATGG 1020
      R  W  Q  D  L  K  R  P  V  D  G  L  N  V  T  I  D  D  M  E
1021  AACGTTACAGGAGGAACATTGAAGAGGCTATTGCTACTGGTAACGTTATATTGCCAGACA 1080
      R  Y  R  R  N  I  E  E  A  I  A  T  G  N  V  I  L  P  D  K
1081  AATCTACCAAAAAGTTGGATATCGATATGCTCGGCAACATGATGGAAGCAAGCGTCCTGT 1140
      S  T  K  K  L  D  I  D  M  L  G  N  M  M  E  A  S  V  L  S
1141  CACCCAACCGCGATTTGTATGGCTCCATCCACAACAACATGCACAGCTTCAGCGCGTACA 1200
      P  N  R  D  L  Y  G  S  I  H  N  N  M  H  S  F  S  A  Y  M
1201  TGCACGACCCCGAACATCGATACCTCGAATCATTCGGTGTGATCGCTGATGAAGCGACGA 1260
      H  D  P  E  H  R  Y  L  E  S  F  G  V  I  A  D  E  A  T  T
1261  CGATGCGCGATCCGTTCTTCTACCGCGTGCATGCGTGGGTCGATGACATCTTCCAGAGTT 1320
      M  R  D  P  F  F  Y  R  V  H  A  W  V  D  D  I  F  Q  S  F
1321  TCAAAGAGGCCCCTCATAACGTGCGCCCATACAGTCGCTCTCAGCTCGAGAACCCTGGCG 1380
      K  E  A  P  H  N  V  R  P  Y  S  R  S  Q  L  E  N  P  G  V
1381  TTCAAGTGACATCCGTCGCGGTGGAGTCTGCTGGCGGCCAACAGAACGTGCTGAACACTT 1440
      Q  V  T  S  V  A  V  E  S  A  G  G  Q  Q  N  V  L  N  T  F
1441  TCTGGATGCAGAGCGATGTGAACCTCTCTAAAGGTTTAGATTTCTCGGACCGCGGGCCGG 1500
      W  M  Q  S  D  V  N  L  S  K  G  L  D  F  S  D  R  G  P  V
1501  TGTACGCGCGCTTCACTCATCTCAACCACAGACCTTTCCGTTACGTTATAAAGGCGAACA 1560
      Y  A  R  F  T  H  L  N  H  R  P  F  R  Y  V  I  K  A  N  N
1561  ACACTGCGTCCGCTCGCCGCACGACCGTGCGCATCTTCATAGCCCCGAAGACAGACGAGC 1620
      T  A  S  A  R  R  T  T  V  R  I  F  I  A  P  K  T  D  E  R
1621  GCAACCTGCCGTGGGCTCTGTCCGACCAACGCAAGATGTTCATTGAGATGGATAGATTCG 1680
      N  L  P  W  A  L  S  D  Q  R  K  M  F  I  E  M  D  R  F  V
1681  TAGTACCCTTGAGCGCTGGCGAAAACACAATTACTCGTCAGTCCACAGAATCGTCACTGA 1740
      V  P  L  S  A  G  E  N  T  I  T  R  Q  S  T  E  S  S  L  T
1741  CAATTCCGTTCGAACAGACGTTCCGCGACCTTTCCATCCAGGGAAGTGACCCTAGGCGTT 1800
      I  P  F  E  Q  T  F  R  D  L  S  I  Q  G  S  D  P  R  R  S
1801  CGGAACTCGCCGCATTTAATTACTGCGGGTGTGGCTGGCCGCAGCACATGCTCGTGCCTA 1860
      E  L  A  A  F  N  Y  C  G  C  G  W  P  Q  H  M  L  V  P  K
1861  AGGGAACGGTCGGCGGTGTTGCCTATCAACTCTTTGTAATGCTTTCCAACTACGAACTGG 1920
      G  T  V  G  G  V  A  Y  Q  L  F  V  M  L  S  N  Y  E  L  D
1921  ATAAGATTGAGCAACCAGACGGCAGGGAACTGAGTTGCGTCGAAGCTTCCATGTTCTGTG 1980
      K  I  E  Q  P  D  G  R  E  L  S  C  V  E  A  S  M  F  C  G
1981  GCTTGAAGGATAAGAAGTACCCAGACGCACGTCCGATGGGCTATCCGTTCGACCGGCCAT 2040
      L  K  D  K  K  Y  P  D  A  R  P  M  G  Y  P  F  D  R  P  S
2041  CCAACTCTGCTACCAACATCGAAGACTTCAGCGCGATGTCCAACATGGGACTGCAAGACA 2100
      N  S  A  T  N  I  E  D  F  S  A  M  S  N  M  G  L  Q  D  I
2101  TAGTCATCAAACTATCCGATGTAACCGAACCAAACCCTAGAAACCCACCCGCTTAACTA 2160
      V  I  K  L  S  D  V  T  E  P  N  P  R  N  P  P  A
2161  CCAAGCTTGCTAAATCATTGTGAATTTACTATCACAAAGGATTTTTGTTTATAATGATAC 2220
2221  TTTATGATACTTAAGTATTACATTTATAATTGAATTTTAAGCTTGAATTATCCTCATAGT 2280
2281  GTTTTTCTTTTAATTTTACTTTTCTTTAAAAATAAAAAATGAAAGGCTGATAGTGTTTAT 2340
2341  CTTGTAAATTCTTAAGTGTAAATATGTTATATATGTTTGTAGAGTTATGTAACTATTATT 2400
2401  GACCCATTTTATAAGGAATAAGACAATAAATAAAACAATGCATTTATAAAAAAAAAAAA 2460
2461  AAAAA 2465
```

FIG. 5. cDNA sequence and deduced protein sequence of *M. sexta* pro-PO. Translation start site is shown in boldface type; polyadenylylation signal and poly(A) tail are underlined. Two peptide sequences obtained by Edman degradation are doubly underlined.

ing the N-terminal methionine is enclosed within a sequence (shown in boldface type) that is ideal for initiation of translation by eukaryotic ribosomes (36). The sequences of two peptides determined by Edman degradation are doubly underlined. The cDNA codes for a protein of 695 amino acids with a M_r of 80,020. There is no evidence of a signal sequence.

Deduced Protein Sequence. The sequence of *M. sexta* pro-PO shows a high degree of identity with other arthropod pro-PO sequences (Fig. 6), especially the *B. mori* pro-PO-2 (37) (77% identity). Identical residues shared with the other sequences are *B. mori* pro-PO-1 (37), 48%; *Drosophila melanogaster* pro-PO (38), 44%; *P. leniusculus* pro-PO (35), 39%. In common with these sequences, the *M. sexta* protein shows similarity to arthropod hemocyanins and larval storage proteins but resembles vertebrate tyrosinases only in the sequence of the copper-binding sites. Given the fact that the insect enzymes do not resemble the vertebrate tyrosinases in their structures and that they participate in a wider variety of reactions (particularly those of sclerotization), the retention of the name PO for the insect enzyme is fully justified.

Particular areas of conserved sequence worthy of note are the Arg-Phe bond at positions 64 and 65 (indicated by arrow), which is cleaved in activation of the *B. mori* pro-POs (37). The N-terminal of the crayfish active PO begins at position 188, resulting from cleavage of the Arg-Thr bond (indicated by arrow) (35). In the other four sequences there is an arginine at position 183 (arrow follows). The sequences after this arginine are similar to the N-terminal PO sequence reported for the housefly *Musca domestica* (39), which is EEATV-VPDG. This result suggests that in the housefly, an Arg-Glu bond in pro-PO may be cleaved in the activation process. We have no information about the N-terminal sequence of the active PO of *M. sexta.* Perhaps activation of pro-PO occurs by cleavage of more than one peptide bond. Taken together with the detergent activation, these facts suggest that the activation of pro-PO involves a conformational change in the polypeptide that can be effected either by proteolytic cleavage at one or more sites or by binding of detergent molecules.

The *M. sexta* pro-PO cDNA was found by using a degenerate oligonucleotide that coded for the conserved sequence GCGEQNM, found in α-macroglobulins and the C3 and C4 complement proteins. In these proteins, the sulfur of the cysteine in this sequence displaces the NH_2 of the glutamine three residues away to give a cyclic thiol ester. When the proteins undergo proteolysis, a conformational change in the polypeptide exposes the thiol ester to the aqueous environment and also activates it so that it reacts avidly with nucleophiles. In the case of the complement proteins, this results in immobilization on cell surfaces.

In the pro-PO sequence there is a region with similarity to the thiol ester sequence of the complement proteins, but the match is not exact. For the five pro-PO sequences beginning at position 618, the corresponding sequences are doubly underlined. This sequence represents the region of the *M. sexta*

```
          1                                                                                                  100
MSPPO     .......... ...MADIFDS FELLYDRPGE PMINTKGEDK VLFELTEQFL TPEYANNGLE LNNRFGDEEE VSRKIILKNL DKIPEFPKAK QLPNDADFSL
BMPPO2    .......... ...MADVFES LELLFDRPNE PLITPKGENN SVFQLTEQFL TEDYANNGIE LNNRFGDDA. .SEKIPLKNL SKLPEFKIAT QLPKDAEFSL
BMPPO1    .......... ...MSDAKNN LLLFFDRPSE PCFMQKGEEN AVFEIPDNYY PEKYQRVSNA IGNRFGSDA. .GRMIPIRNI A.LPNLDLPM ELPYNEQFSL
DMPPO     .......... ..MTNTDLKA LELMFQRPLE PAFTTRDSGK TVLELPDSFY TDRYRNDTEE VGNRFSKDVD L..KIPIQEL SNVPSLEFTK KIGLKNQFSL
PLPPO     MQVTQKLLRR DTEMADAQKQ LLYLFERPYD PINAPRADGS FLYAVAG... .......ADT VATRFGVAPT STVTVPARPD ADRRLLGRAP SVPRGAVFSF

          101                                                                                                200
MSPPO     FLPSHQEMAN EVIDVLMSVT ENQLQELLST CVYARINLNP QLFNYCYTVA IMHRRDTGKV RVQNYAEIFP AKFLDSQVFT QAR.EAAAVI PKTIPRTPII
BMPPO2    FLPKHQEMAN ELLGVLMDVP ENELQDLLST CAFARVNLNP QLFNYCYSVA LMHRRDTRKV RVKNFAEVFP SKFLDSQVFT QAR.ETAAVI PPDVPRIPII
BMPPO1    FVPKHRKLAG RLIDIFMGMR D..VEDLQSV CSYCQLRINP YMFNYCLSVA ILHRPDTKGL SIPTFAESFP DKFMDPKVFR QAR.EVSSVV PSG.ARMPIV
DMPPO     FNNRHREIAS ELITLFMSAP N..LRQFVSL SVYTKDRVNP VLFQYAYAVA VAHRPDTREV PITNISQIFP SNFVEPSAFR DAR.QEASVI GESGARVHVD
PLPPO     FIRSHREAAR DLCDVLMKTQ NST..DLMQL AASVRRHVNE NLFIYALSFT ILRKQELRGV RLPPILEVFP HKFIPMEDLT SMQVEVNRTP PTATTPLVIE

          201                                                                                                300
MSPPO     IPRDYTATDL EEEHRLAYWR EDLGINLHHW HWHLVYPFSA SDEKIVAKDR RGELFFYMHQ QIIARYNCER LCNSLKRVKK FSDWREPIPE AYYPKLDSLT
BMPPO2    IPRDYTATDL EEEHRLAYWR EDIGINLHHY HWHLVYPFTA NDLSIVAKDR RGELFFYMHQ QVIARFNCER LCNSLKRVKK FSNWREPIPE AYFPKLDSLT
BMPPO1    IPSNYTASDT EPEQRVAYFR EDIGINLHHW HWHLVYPFDA ADRAIVNKDR RGELFYYMHQ QIIARYNVER MCNNLSRVRR YNNFRAAIEE GYFPKLDSTV
DMOPPO    IPQNYTASDR EDEQRLAYFR EDIGVNSHHW HWHLVYP.TT GPTEVVNKDR RGELFYYMHH QILARYNVER FCNNLKKVQP LNNLRVEVPE GYFPKILSST
PLPPO     YGPEFANTNQ KAEHRVSYWR EDFGINSHHW HWHLVYPIEM N....VNRDR KGELFYYMHQ QMVARYDWER LSVNLNRVEK LENWRVPIPD GYFSKLTANN

          301                                                                                                400
MASPPO    SARGWPPRQA GMRWQDLKRP VDGL.NVTID DMERYRRNIE EAIATGNVIL PDKSTKKLD. ........ID MLGNMMEAS. VLSPNRDLYG SIHNNMHSFS
BMPPO2    SSRGWPPRQS GMQWQDLNRA AEGL.FVTID EMERWRRNVE EAIATGTVRL PNGQTRPLD. ........ID TLGNMLESS. ALSPNRELYG SIHNNGHSFT
BMPPO1    ASRAWPPRFA GTTIRDLDRP VDQI.RSDVS ELETWRDRFL QAIENMSVML PNGRQLPLDE E.....TGID VLGNLMESS. IISRNRPYYG DLHNMGHVFI
DMOPPO    NNRTYPARVT NQKLRDVDRH DG...RVEIS DVERWRDRVL AAIDQGYVED SSGNRIPLDE V.....RGID ILGNMIEASP VLSINYNFYG NLHNEGHNII
PLPPO     SGRPWGTRQD NTFIKDFRRN DAGLDFIDIS DMEIWRSRLM DAIHQGYMLN RNGERVPLSD NVTTGKRGID ILGDAFEADA QLSPNYLFYG DLHNTGHVLL

          401                                                                                                500
MASPPO    AYMHDPEHRY LESFGVIADE ATTMRDPFFY RVHAWVDDIF QSFKEAPHNV RPYSRSQLEN PGVQVTSVA. .VESAGGQQN VLNTFWMQSD VNLSKGLDFS
BMPPO2    AYMHDPEHRY LEQFGVIADE ATTMRDPFFY RWHAYIDDVF QKHKESAY.V RPYTRSELEN PGVQVRSVS. .VETPGGQPN TLNTFWMLSD VNLSRGLDFS
BMPPO1    SYSHDPDHRH LEQFGVMGDS ATAMRDPVFY RWHAYIDDIF HLYK...YKL TPYGNDRLDF PNIRVSSVS. .IE.GGGTPN TLNTLWEQST VDLGRGMDFT
DLPPO     SFAHDPDYRH LEDFGVMGDV TTAMRDPIFY RWHGFIDTVF NKFKT...RL NPYNAGELNF DGITVDYIEA KIGKSNTKAN TLLTYWQKSS ADLAAGLDFG
PLPPO     AFCHDNDNSH REEIGVMGDS ATALRDPVFY RWHKFVDDIF QEYKLTQ... PPYTMEDLSL PGVVLDKVGV ...VRNDQLN TLTTGWSVRE FEASRGLDFN

          501                                                                                                600
MSPPO     DRGP..VYAR FTHLNHRPFR YVIKANNT.A SARRTTVRIF IAPKTDERNL PWALSDQRKM FIEMDRFVVP LSAGENTITR QSTESSLTIP FEQTFRDLSI
BMPPO2    DNGP..VYAR FTHLNYRHFS YRINVNNT.G SSRRTTVRIF ITPKFDERNV PWIFSDQRKM CIEMDRFVTV LNAGENNIVR QSTESSITIP FEQTFRDLSA
BMPPO1    PRGS..VLAR FTHLQHDEYN YVIEVNNTGG SSVMGMFRIF IAPTVDESGK PFSFDEQRKL MIELDKFSQG VKPGNNTIRR KSIDSSVTIP YERTFRNQAD
DMPPO     PTTDRNIFAS FTHLQNAPFT YTFNVTNN.G ARRTGTCRIF ICPKVDERNQ ALNLEEQRLL AIEMDKFTVD LVPGENTIRR QSTESSVAIP FERSFRPVGA
PLPPO     SPNP..VTAH YPS.RPCTLH LPSPDNKQHR KPKSVTVRIY MAPKHNERGL EMGFMEQRLL WAEMDKFTQD LKPGQNQIVR ASNLSSITNP SGYTFRSLEA

          601                                                                                                700
MSPPO     .QGSDPRRSE LAAFNYCGCG WPQHMLVPKG TVGGVAYQLF VMLSNYELDK IEQPDGRELS CVEASMFCGL KDKKYPDARP MGYPFDRPSN S......ATN
BMPPO2    .QGNDPRRDE LATFNYCGCG WPQHMLVPKG TEAGMPFQLF VMLSNYDLDR IDQDDGKQLT CVEASSFCGL KDKKYPDRRA MGFPFDRPSS S......ATS
BMPPO1    .RPADPGTAG AAEFDFCGCG WPHHMLVPKG TTQGYPMVLF VMVSNWNDDR VEQDLVG..S CNDAASYCGI RDRKYPDRRA MGFPFDRPAP A......ATT
DMPPO     .DYQPKAADE LARFKFCGCG WPQHLLLPKG NAQGMLFDLF VMISDYSQDS VEQPKTPNDA CSTAYSFCGL KDKLYPDRRT MGYPFDRRLP N......ANL
PLPPO     VNPANPGPPA NAETNFCGCG WPEHLLLPRG KPEGMTYQLF FMLTDLEKDQ VDQPAGPR.R CANAVSFCGI LDSKFPDKRP MGFPFDRRPP PRLQDAEVTS

          701                          734
MSPPO     IEDFSAMSNM GLQDIVIKLS DVTEPNPRNP PA..
BMPO2     LQDF.ILPNM GLQDITIQLQ NVTEPNPRNP PMSV
BMPO1     LSDF.LRPNM AVRDCIVRFT DRTRQRGQQG ....
DMPPO     TELVGAFGNM AKTDLRIVFN DRVIDKA... ....
PLPPO     VADYARLSNM TVQDITITFL TTASRSRHDG PI..
```

FIG. 6. Alignment of five arthropod pro-PO sequences (alignment with PILEUP program, Genetics Computer Group, Madison, WI) MSPPO, pro-PO of *M. sexta*; BMPPO2, pro-PO2 of *B. mori* (37); BMPPO1, pro-PO1 of *B. mori* (37); DMPPO, pro-PO of *D. melanogaster* (38); and PLPPO, pro-PO of *P. leniusculus* (35). Underlined regions are the copper-binding sites. The region at which the degenerate oligonucleotide annealed to the *M. sexta* cDNA in the initial PCR is shown with double underlining. This oligonucleotide was designed to match the thiol ester region of α-macroglobulin and complement proteins C3 and C4. Arrows show potential cleavage sites for proteolytic activation of the zymogen (see text).

cDNA where the original PCR primer annealed. It can be seen that the first three residues correspond exactly to the thiol ester region of the complement proteins, but the glutamine appears in only three of the sequences, and it is displaced by one residue. Similar sequences are present in arthropod hemocyanins (40). Resemblance of this sequence to those in α-macroglobulin or the complement proteins may be purely coincidental, and we do not know whether these residues actually participate in thiol ester formation. If they do, however, the formation of a reactive thiol ester might participate in the physiological function of PO to bind to foreign surfaces, and it might explain some of those properties of the active enzyme that have frustrated investigators. It would also relate PO, a component of the arthropod immune system, with complement, a component of the vertebrate immune system.

We are grateful to Drs. Masaaki Ashida and Eiji Ohnishi for sharing their data with us before publication, and to Dr. Michael R. Kanost, who supplied the hemocyte library, read the manuscript, and made valuable suggestions for improving it. We thank Kyril Choumarov, Dianzheng Zhang, and Joy J. Winzerling for valuable help in completing this work and the Harvard Microsequencing Laboratory for protein sequence analysis. This research was supported by National Institutes of Health Grants AI14753 (to M.S.) and AI32595 (to J.H.L.), by the John D. and Catherine T. MacArthur Foundation (no. 8900408 to M.H.), by the Swedish Natural Science Research Council, and by the Swedish Research Council for Forestry and Agricultural Research (to K.S.).

1. Prota, G. (1992) *Melanins and Melanogenesis* (Academic, San Diego).
2. Hearing, V. J. & Tsukamoto, K. (1991) *FASEB J.* **5,** 2902–2909.
3. Sugumaran, M. (1991) *FEBS Lett.* **293,** 4–10.
4. Sugumaran, M. (1988) *Adv. Insect Physiol.* **21,** 179–231.
5. Sugumaran, M. (1991) in *The Physiology of the Insect Epidermis*, eds. Binnington, K. & Retnakaran, A. (Commw. Sci. Ind. Res. Org. Press, Canberra, Australia), pp. 141–168.
6. Thompson, D. C., Thompson, J. A., Sugumaran, M. & Moldeus, P. (1993) *Chem. Biol. Interact.* **86,** 129–162.
7. Andersen, S. O. (1991) in *The Physiology of the Insect Epidermis*, eds. Binnington, K. & Retnakaran, A. (Commw. Sci. Ind. Res. Org. Press, Canberra, Australia), pp. 123–140.
8. Hopkins, T. L. & Kramer, K. J. (1992) *Annu. Rev. Entomol.* **37,** 273–302.
9. Saul, S. J. & Sugumaran, M. (1988) *FEBS Lett.* **237,** 155–158.
10. Saul, S. J. & Sugumaran, M. (1989) *FEBS Lett.* **249,** 155–158.
11. Saul, S. J. & Sugumaran, M. (1990) *J. Biol. Chem.* **265,** 16992–16999.
12. Saul, S. J., Dali, H. & Sugumaran, M. (1991) *Arch. Insect Biochem. Physiol.* **16,** 123–137.
13. Sugumaran, M., Giglio, L. B., Kundzicz, H., Saul, S. J. & Semensi, V. (1992) *Arch. Insect Biochem.* **19,** 271–283.
14. Saul, S. J. & Sugumaran, M. (1989) *FEBS Lett.* **255,** 340–344.
15. Sugumaran, M., Semensi, V., Kalyanaraman, B., Bruce, J. M. & Land, E. J. (1992) *J. Biol. Chem.* **267,** 10355–10361.
16. Ricketts, D. & Sugumaran, M. (1994) *J. Biol. Chem.* **269,** 22217–22221.
17. Sugumaran, M. & Kanost, M. (1993) in *Parasites and Pathogens of Insects*, eds. Beckage, N. E., Thompson, S. N. & Federici, B. A. (Academic, San Diego), pp. 317–342.
18. Ashida, M. & Yamazaki, H. L. (1990) in *Molting and Metamorphosis*, eds. Ohnishi, E. & Ishizaki, H. (Japan Sci. Press, Tokyo), pp. 239–265.
19. Söderhäll, K. (1982) *Dev. Comp. Immunol.* **6,** 601–611.
20. Lai-Fook, J. (1966) *J. Insect Physiol.* **12,** 195–226.
21. Ashida, M. (1971) *Arch. Biochem. Biophys.* **144,** 749–762.
22. Aso, Y., Kramer, K. J., Hopkins, T. L. & Lookhart, G. L. (1985) *Insect Biochem.* **15,** 9–17.
23. Andersson, K., Sun, S. C., Boman, H. G. & Steiner, H. (1989) *Insect Biochem.* **19,** 629–637.
24. Pau, R. N. & Eagles, A. M. (1975) *Biochem. J.* **149,** 707–712.
25. Tsukamoto, T., Ishiguro, M. & Funatsu, M. (1986) *Insect Biochem.* **16,** 573–581.
26. Durrant, H. J., Ratcliffe, N. A., Hipkin, C. R., Aspán, A. & Söderhäll, K. (1993) *Biochem. J.* **289,** 87–91.
27. Heyneman, R. A. (1965) *Biochem. Biophys. Res. Commun.* **21,** 162–169.
28. Aspán, A. & Söderhäll, K. (1991) *Insect Biochem.* **21,** 363–373.
29. Bell, R. A. & Joachim, F. G. (1976) *Ann. Entomol. Soc. Am.* **69,** 365–373.
30. Sottrup-Jensen, L. (1987) in *The Plasma Proteins*, ed. Putnam, F. W. (Academic, Orlando, FL), Vol. 5, pp. 191–291.
31. Spycher, S. E., Arya, S., Isenman, D. E. & Painter, R. H. (1987) *J. Biol. Chem.* **262,** 14606–14611.
32. Hall, M., Söderhäll, K. & Sottrup-Jensen, L. (1989) *FEBS Lett.* **254,** 111–114.
33. Enghild, J. J., Thøgersen, I. B., Salvesen, G., Fey, G. H., Figler, N. L., Gonias, S. L. & Pizzo, S. V. (1990) *Biochemistry* **29,** 10070–10080.
34. Thøgersen, I. B., Salvesen, G., Brucato, F. H., Pizzo, S. V. & Enghild, J. J. (1992) *Biochem. J.* **285,** 521–527.
35. Aspán, A., Huang, T.-S., Cerenius, L. & Söderhäll, K. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 939–943.
36. Kozak, M. (1986) *Cell* **44,** 283–292.
37. Kawabata, T., Yasuhara, Y., Ochiai, M., Matsuura, S. & Ashida, M. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 7774–7778.
38. Fujimoto, K., Okino, N., Kawabata, S., Iwanaga, S. & Ohnishi, E. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 7769–7773.
39. Hara, T., Miyoshi, T. & Tsukamoto, T. (1993) *Comp. Biochem. Physiol. B.* **106,** 287–292.
40. Beintema, J. J., Stam, W. T., Hazes, B. & Smidt, M. P. (1994) *Mol. Biol. Evol.* **11,** 493–503.